

5

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-305588

(43)Date of publication of application : 02.11.2000

---

(51)Int.Cl. G10L 15/00  
G10L 13/00  
G10L 15/10  
G10L 15/28  
G10L 19/02  
G10L 19/00  
H04N 5/765  
H04N 5/781  
H04N 5/91

---

(21)Application number : 11-115981

(71)Applicant : HITACHI LTD

(22)Date of filing : 23.04.1999

(72)Inventor : IMAI TSUTOMU  
OWASHI HITOAKI

---

### (54) USER DATA ADDING DEVICE AND USER DATA REPRODUCING DEVICE

#### (57)Abstract:

PROBLEM TO BE SOLVED: To provide a user data adding device and a user data reproducing device facilitating a management of music data and user data and processing the user data synchronously with the music data to display on a monitor or the like at any time.

SOLUTION: A user data adding device 111 embedding user data such as a music name, a singer name, words, a CD jacket and a static image data of the singer into music data without separating from the user data and a user data reproducing device are provided. This constitution displays a static image and character data during reproducing of the music data and displays the user data continuously synchronizing with the music.

---

### CLAIMS

---

[Claim(s)]

[Claim 1] User data additional equipment comprising:

An image input terminal which inputs a picture.

A music input terminal which inputs music.

A text input terminal which inputs a text.

A still picture encoder which compresses said inputted picture and a music encoder which compresses said inputted music. A switch which distributes said compressed music stream to a still picture adjunct, a text adjunct, or a synchronous text adjunct. A voice recognition part which recognizes a sound from said inputted music and is made into text data. A still picture adjunct which adds still picture data compressed into said compressed music stream. A text adjunct which adds text data inputted into said compressed music stream from a text input terminal and a synchronous text adjunct which adds data text-ized by said compressed music stream by a voice recognition part.

[Claim 2] User data additional equipment comprising:

A switch which distributes a compressed music stream to a still picture adjunct, a text adjunct, or a synchronous text adjunct.

A music decoder which elongates said compressed music stream.

A voice recognition part which recognizes a sound from a music encoder which represses said elongated music data and said elongated music data and is made into text data.

A still picture adjunct which adds still picture data compressed into said compressed music stream and a text adjunct which adds text data to said compressed music stream and a synchronous text adjunct which adds data text-ized by said compressed music stream by a voice recognition part.

[Claim 3] User data playback equipment comprising:

A music decoder which elongates a compressed music stream.

A user data judgment part which judges whether data other than music data of said elongated music stream exists.

A switch which sends user data to a still picture decoder or a text display area.

Said still picture decoder which elongates said compressed still picture data and said text display area which processes user data of a text, generating picture terminal which outputs said still picture and said text and a music output terminal which outputs said music data.

[Claim 4] User data additional equipment wherein said still picture adjunct embeds and generates still picture data as user data to a generate time of a music stream in claim 1 statement.

[Claim 5] User data additional equipment wherein alphabetic data embeds and generates said text adjunct as user data to a generate time of a music stream in claim 1 statement.

[Claim 6] User data additional equipment wherein alphabetic data which synchronized with a generate time of a music stream as user data at music embeds and generates said synchronous text adjunct in claim 1 statement.

[Claim 7] In claim 1 statement said still picture adjunct or a text adjunct User data additional equipment wherein AAU (Audio Access Unit audio access unit) of a head of music of a music stream generates a music stream which comprised non-sound data and user data.

[Claim 8] In claim 1 statement said still picture adjunct or a text adjunct User data additional equipment wherein some AAU(s) (Audio Access Unit audio access unit) generate a music stream which comprised non-sound data and user data from a head of music of a music stream.

[Claim 9] In claim 1 statement said still picture adjunct or a text adjunct User data additional equipment wherein some AAU(s) (Audio Access Unit audio access unit) of a music stream generate a music stream which comprised music data and user data.

[Claim 10] In claim 1 statement said synchronous text adjunct User data additional equipment wherein all the AAU(s) (Audio Access Unit audio access unit) of a music stream generate a music stream which comprised music data and user data.

[Claim 11] User data additional equipment with which said music encoder is characterized by a thing in 1 AAU (Audio Access Unit audio access unit) for which it leaves a field of user data beforehand and compression processing is performed in claim 1 statement.

[Claim 12] User data additional equipment wherein said still picture adjunct embeds and generates still picture data as user data to a generate time of a music stream in claim 2 statement.

[Claim 13] User data additional equipment wherein alphabetic data which synchronized with a generate time of a music stream as user data at music embeds and generates said synchronous text adjunct in claim 2 statement.

[Claim 14] User data additional equipment wherein said synchronous text adjunct embeds and generates alphabetic data which synchronized with a generate time of a music stream as user data at music in claim 2 statement.

[Claim 15] In claim 2 statement said still picture adjunct or a text adjunct User data additional equipment wherein AAU (Audio Access Unit audio access unit) of a head of music of a music stream generates a music stream which comprised non-sound data and user data.

[Claim 16] In claim 2 statement said still picture adjunct or a text adjunct User data additional equipment wherein some AAU(s) (Audio Access Unit audio access unit) generate a music stream which comprised non-sound data and user data from a head of music of a music stream.

[Claim 17] In claim 2 statement said still picture adjunct or a text adjunct User data additional equipment wherein some AAU(s) (Audio Access Unit audio access unit) of a music stream generate a music stream which comprised music data and user data.

[Claim 18]In claim 2 statementsaid synchronous text adjunctUser data additional equipmentwherein all the AAU(s) (Audio Access Unitaudio access unit) of a music stream generate a music stream which comprised music data and user data.

[Claim 19]In claim 2 statementsaid still picture adjunct or a text adjunctUser data additional equipmentwherein AAU (Audio Access Unitaudio access unit) of the last of music of a music stream generates a music stream which comprised non-sound data and user data.

[Claim 20]In claim 2 statementsaid still picture adjunct or a text adjunctUser data additional equipmentwherein some AAU(s) (Audio Access Unitaudio access unit) generate a music stream which comprised non-sound data and user data from the last of music of a music stream.

[Claim 21]User data additional equipmentwherein said music encoder leaves a field of user data beforehand in AAU (Audio Access Unitaudio access unit) and performs compression processing in claim 2 statement.

[Claim 22]In claim 3 statementsaid user data judgment partUser data playback equipment 0 or in all data to the next AAU (Audio Access Unitaudio access unit) following non-sound data or music data being altogether except onewherein it processes by regarding as user data.

[Claim 23]User data playback equipment when AAU (Audio AccessUnitaudio access unit) is non-sound data in claim 3 statementwherein said user data judgment part processes by regarding data to the next AAU as user data.

[Claim 24]In claim 22 or claim 23 statementsaid user data judgment partUser data playback equipment searching a header of a compressed still picture in user datajudging with still picture data which had user data compressed when a header is foundand judging with text data when a header is not found.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention]This invention relates to the device which plays the music stream downloaded by the music distribution of the network which uses the music distribution which uses the device and communications satellite which carry out digital compression of the music and are recorded on a recording mediuma personal computeretc.It is related with the user data additional equipment and user data playback equipment which add text data and still picture data to the music stream compressed especially.

[0002]

[Description of the Prior Art]As an example which carries out the conventional digital compression and performs a music distributionthe music distribution which uses a

communications satellite and the music distribution downloaded from the homepage on the Internet using a personal computer occur.

[0003] sky perfect PC (SKY PerfecPC!) "" [ which is data broadcasting for PC ] Be as what performs a music distribution using a communications satellite. This performs the data of a homepage program download of music data etc. using a communications satellite.

[0004] A fixed fee is paid the music data on Web by which digital compression was carried out as what performs the music distribution which on the other hand uses the Internet which uses a personal computer etc. is downloaded and it reproduces with the personal computer etc. which can elongate the music data.

[0005] There is a music compression equipment which carries out digital compression of the music which is not compressed in real time.

[0006]

[Problem(s) to be Solved by the Invention] In the conventional music distribution it is download of only a music stream and a track name a singer name etc. which are user data were attached to the file name recorded on recording media such as a hard disk and a memory. Although the information on a track name a singer name still picture data etc. may also be downloadable as user data apart from music data it is necessary to make a file name the same or to record the music data and user data by which digital compression was carried out in that case on the same directory. However for example it processed music data and user data by taking a synchronization it was difficult to display words at any time according to music.

[0007] When music was recorded in real time it is only music and there was no user data.

[0008]

[Means for Solving the Problem] In order to solve an aforementioned problem in this invention a device and user data playback equipment which are embedded at music data are provided without making user data such as a track name a singer name words CD jackets and a singer's still picture data different from music data. It becomes possible to display user data while playing music data or to make it synchronize with music and to display user data continuously by this.

[0009]

[Embodiment of the Invention] Hereafter the example of this invention is described using a drawing.

[0010] As user data a track name and a singer name words Compressing music into real time as a device which adds CD jackets and a singer's still picture data. How to add user data to the music stream which once recorded the case where user data is added and the distributed music stream on recording media such as a hard disk and a memory is explained.

[0011] The case where user data is added is explained as the first example compressing the music from a microphone or an external device into real time.

It explains using drawing 1 and drawing 16 about the case of MPEG Audio LayerII as musical digital compression technology.

[0012]In drawing 11 101 an image input terminal and 102 a music input terminal and 103 A text input terminal 104 -- a still picture encoder and 105 -- a music encoder and 106 -- as for a text adjunct and 110a voice recognition part and 108 are [ the user data additional equipment and 112 ] recording mediasuch as a hard disk and a memorya synchronous text adjunct and 111 a still picture adjunct and 109 a switch and 107.

[0013]Imaging devicesuch as a cameraare connected to the image input terminal 101a microphone or a music reproduction device is connected to the audio input terminal 102and it is connected with a keyboard etc. at the text input terminal 103. Digital compression is carried out by the music encoder 105and the music incorporated from the music input terminal 102 generates a music stream. Nextaccording to the user data to addit changes to the still picture adjunct 108the text adjunct 109and the synchronous text adjunct 110 with the switch 106. The picture inputted into the music stream which the music encoder 105 generated with the image input terminal 101 in the still picture adjunct 108 with the still picture encoder 104 JPEG (Joint PhotographicCoding Experts Group)The still picture picture compressed into the format of GIF (Graphics Interchange Format) etc. is added. In the text adjunct 109the text data inputted into the music stream which the music encoder 105 generated by the keyboard etc. from the text input terminal 103 is added. From the music data inputted into the music stream which the music encoder 105 generated from the music input terminal 102by the voice recognition part 107from speech informationwords are recognizedand it text-izesand adds at the synchronous text adjunct 110.

[0014]Herethe music stream which the music encoder 105 generates is explained using drawing 15.

[0015]In drawing 15 1501 is the minimum unit which can be independently decoded to a music signal one by one by AAU (Audio Access Unitaudio access unit) which constitutes the music stream. AAU1501 consists of the header 1502the bit allocation 1503SCFSI(Scalefactor Selection Informationscale-factor selection information)1504the scale factor 1505and the sample 1506.

[0016]The music stream which the text adjunct 108 in the case of adding the text of a track name and a singer name with little data volume as the first example as user data generates is explained using drawing 5.

[0017]As for user data and 503in drawing 55 501 is [ music data and 505 ] non-sound data and 506 user data music data and 504 non-sound data and 502. From the non-sound data of 501 to the music data of 504 is a music stream for one musicand the non-sound data of 505 or subsequent ones shows the music stream of the following music. In this casethe case where constitute AAU of the head of two or more AAU(s) which constitute each music stream from the non-sound data 501 and the user data

502namelythe user data 502 is added to the non-sound data 501 is shown.

[0018]Herethe method of creation of the non-sound data 501 is explained. In the case of MPEG AudioLayerII48 kHz of sampling frequencyand bit rate 256k bpsand a stereo1AAU is 6144 bits. Non-sound data can be used as non-sound data only at 120 bits of a header (32 bits) and a bit allocation (88 bits) by setting 32 bits of the header 1602 to 0xFFFD C40Cand setting all of 88 bits of the bit allocation 1603 to 0. That isit becomes possible for 1AAU to remain and to assign  $6144 - 120 = 6024$  bit to user data.

[0019]Nexthow to take out user data is explained using drawing 16. In drawing 161601 a recording medium and 1602 a music decoder and 1603 A user data judgment partAs for a text display area and 1607a switch and 1605 are [ a generating picture terminal and 1609 ] music output terminals user data playback equipment and 1608 a still picture decoder and 1606 1604.

[0020]The music data stream with user data read from the recording medium 1601 is decoded by the music decoder 1602and is outputted from the loudspeaker connected to the music output terminal 1609.

[0021]In the user data judgment part 1603the data from data after music decoding was carried out to the header of the next AAU is investigated. Usuallycalculating so that most 1AAU may be filledmusic data is compressed and an encoder embeds 0 and 1 altogether into the surplus portion. Thereforeat the time of 0 or 1no user data are altogether outputted from the generating picture terminal 1608 to the next AAU as what not existing. When the header of still picture datasuch as JPEG and GIFis not found by the next AAUa text is displayed on the monitor etc. which regard it as text datadecode a text by the text display area 1606and are connected to the generating picture 1608. Since silent AAU does not existit may usually completely process as AAU in which user data is contained at the time of silent AAU.

[0022]It is assumed that it is the bottom which decodes AAU which comprises non-sound data and user data with the playback equipment whichon the other handdoes not have the function to process user data other than user data playback equipment 1607. In this casethe user data which processes only non-sound data and follows it will be disregarded. Thereforefrom a speakersince a non-sound will be reproduced without being able to hear a soundit does not have influence at all. Since the samples by which the regeneration time of 1AAU is included in 1AAU are 1152 samplesthey are set to  $1/48(\text{kHz}) \times 1152 = 24$  (ms). Thereforewhen it has composition of the music stream of drawing 1it becomes silent data for 24 ms by relation of music and music.

[0023]As the second exampleas user dataJPEG (Joint PhotographicCoding Experts Group)It is considered as CD jackets compressed by GIF (Graphics Interchange Format) etc.or the still picture data for one sheet of a singerand the music stream which a still picture adjunct when the size is comparatively small generates is explained using drawing 6.

[0024]As for non-sound data and 604in drawing 6non-sound data and 602 are [ music data and 606 ] music data user data and 605 user data and 603 601. From the non-

sound data of 601 to the music data of 606 is music data for one music.

[0025]Comparatively little still picture data compressed by JPEGGIFetc. also needs the user data size of several kilobytes – tens of K bytes. Therefore since AAU which comprises non-sound data and user data is insufficient at onesome are prepared so that the amount of still picture data may be suited. As the composition of non-sound data and user data explained above it is usable 6024 bits to user data. For example when using 1 K byte ( $1024 \times 8 = 8192$  bit) of still picture data as user data it corresponds to this by being referred to as two AAU(s) of the composition of non-sound data and user data.

[0026]How to take out user data is explained using drawing 16.

[0027]In drawing 16 the music data stream with user data read from the recording medium 1601 is decoded by the music decoder 1602 and is outputted to the music output terminal 1609. By the user data judgment part 1603 the data from data after music decoding was carried out to the header of the next AAU is investigated like the case of the above-mentioned example 1. At the time of 0 or 1 no user data are altogether outputted from the generating picture terminal 1608 to the next AAU as what not existing. When the header of still picture data such as JPEG and GIF is found by the next AAU output a still picture to the monitor etc. which process the still picture decoder 1605 with the switch 1604 and are connected to the generating picture but. When still picture data is not settled in 1 AAU still picture decoder processing is performed as what has the remaining data also in the next AAU and it outputs from a generating picture terminal.

[0028]The case where they are tens of K bytes – hundreds of K bytes and the still picture data of several sheets whose still picture data size of one sheet is comparatively large are used as user data as the third example is explained using drawing 1 and drawing 7. In this case if it has composition of non-sound data and user data like the composition of AAU shown by drawing 6 since user data size is large AAU of the composition of non-sound data and user data will increase.

Therefore a non-sound is long when this music stream is decoded. So when user data size is large it has composition of the music stream shown in drawing 7.

[0029]in drawing 7 -- 701 -- as for user data and 705 user data and 703 are [ user data and 707 ] music data music data and 706 music data and 704 music data and 702. From the music data of 701 to the music data of 707 is music data for one music. Thus a silent portion can be lost with constituting AAU from music data and user data. However it is necessary to compress music data with the music encoder 105 being conscious of the size beforehand used as user data in this case.

[0030]For example when making 10 K bytes ( $10 \times 1024 \times 8 = 81920$  bit) of still picture data into user data 6044 bits is assigned to the music data 701 among 6144 bits of 1 AAU and the remaining 100 bits are assigned to the user data 702. If such assignment is used it is realizable by being set to  $81920/100 = 819.2$  and considering 820 AAU(s) as the composition of the music data 701 and the user data 702. Since



1AAU is 24 ms it considers the music stream for  $820 \times 24 = 19680$  ms as the composition of music data and user data.

[0031] Since it is the same as the above-mentioned example 2 how to take out user data is omitted.

[0032] The case where take music and a synchronization and words are displayed on a screen as the fourth example is explained using drawing 1 and drawing 8.

[0033] In drawing 8 -- 801 -- music data and 802 -- as for music data and 806 music data and 804 are [ music data and 808 ] user data an user datum and 807 user data and 805 user data and 803. From the music data of 801 to the user datum of 806 is a music stream for one music and the music data of 807 or subsequent ones shows the music data of the following music.

[0034] In drawing 1 the music inputted from the music input terminal 102 recognizes a sound by the voice recognition part 107 and is taken as text data. In the synchronous text adjunct 110 text data is added to the music stream outputted from the music encoder 105 at any time. With the music encoder 105 the music stream which vacated the portion of user data beforehand is created like the example 3.

[0035] How to take out user data is explained using drawing 16. In drawing 16 the music stream with user data read from the recording medium 1601 is decoded by the music decoder 1602 and is outputted from the music output terminal 1609. On the other hand in the user data judgment part 1603 the data from data after music decoding was carried out to the header of the next AAU is investigated. At the time of 0 or 1 no user data are altogether outputted from the generating picture terminal 1608 to the next AAU as what not existing. When the header of still picture data such as JPEG and GIF is not found by the next AAU user data is sent to the text display area 1606 as text data and is outputted from the image output part 1608 at any time.

[0036] How to add user data to the music stream which once recorded the music stream distributed as the second example using the communications satellite and the network on recording media such as a hard disk and a memory is explained. How to download music data is explained to a recording medium using drawing 3 and drawing 4.

[0037] Drawing 3 is a figure explaining signs that it downloads using a communications satellite. As for a parabolic antenna and 302 in drawing 3 a recording medium and 304 are personal computers a receiving board and 303 301. Only the music stream and the music stream of a sample which the music stream of several channels sent from the communications satellite was received by the parabolic antenna 301 and paid the fee on the receiving board 302 are recorded on the recording medium 303. Then the downloaded music stream is elongated with a personal computer etc. and it plays.

[0038] Drawing 4 is a figure explaining signs that the information on the Internet is downloaded via a telephone line etc. In drawing 4 401 shows the Internet 402 shows a modem and the recording medium 404 shows the personal computer 403. In this case music data is registered into Web as a file a fixed fee is paid and it downloads to

the recording medium 403 via the modem 402 using a telephone line. The downloaded music SUTORIMU \*\* is elongated with a personal computer etc. and it plays.

[0039] First the operation which adds user data to the music stream downloaded to the recording medium using drawing 2 is explained. in drawing 2 -- 201 -- a recording medium and 202 -- a switch and 203 -- as for a still picture adjunct and 207 a music encoder and 205 are [ a synchronous text adjunct and 209 ] user data additional equipment a text adjunct and 208 a voice recognition part and 206 a music decoder and 204.

[0040] In the digital compression by the music distribution which uses a communications satellite and a network many MPEG Audio and Layer III(s) (MP3) which are high compression are used from MPEG Audio and Layer II.

[0041] As user data a track name and a singer name words CD jackets and a singer's still picture data can be considered like the first example. The music stream generated using this invention about each case is explained.

[0042] The case where the text data of a track name and a singer name with little data volume as the first example is used as user data is explained using drawing 9 and drawing 10.

[0043] As for non-sound data and 902 in drawing 9 music data and 904 are music data user data and 903 901.

[0044] In drawing 2 the music stream read from the recording medium 201 is sent to the text adjunct 207 by the switch 202 and text data is added to it by the text adjunct 207 and it is recorded on the recording medium 201.

[0045] In making the text data of a track name and a singer name with little data volume into user data as shown in drawing 9 it realizes by constituting AAU of the head of the AAU(s) which constitute a music stream from the non-sound data 901 and the user data 902 namely adding the user data 902 to the non-sound data 901.

[0046] On the other hand since the downloaded audio stream is one file for every music it tends to search AAU of the last (the last of music) of a file. Therefore adding AAU (the non-sound data 1003 user data 1004) which contained user data in the end of a file as shown in drawing 10 can also be realized.

[0047] How to take out user data is explained using drawing 16.

[0048] In drawing 16 the music data stream with user data read from the recording medium 1601 is decoded by the music decoder 1602 and is outputted from the music output terminal 1609. In the user data judgment part 1603 the data from data after music decoding was carried out to the header of the next AAU is investigated. At the time of 0 or 1 no user data are altogether outputted from the generating picture terminal 1608 to the next AAU as what not existing. When the header of still picture data such as JPEG and GIF is not found by the next AAU it is judged as text data processes by the text display area 1606 and outputs from 1608.

[0049] As the second example it is considered as CD jackets compressed as user data by JPEG GIF etc. or the still picture data for one sheet of a singer and the case where

the size is comparatively small is explained using drawing 11 and drawing 12.

[0050]As for user data and 1103in drawing 111101 is [ user data and 1105 ] music data non-sound data and 1104 non-sound data and 1102.

[0051]Even when the compressed still picture data size is smallthe user data size of several kilobytes – tens of K bytes is needed. Therefore since AAU which comprises non-sound data and user data is insufficient at oneit realizes by adding some to the head of a file so that the amount of still picture data may be suited. Considering AAU (the non-sound data 1204user data 1205) which contained user data as shown in drawing 12 as the composition finally added partly can also be realized.

[0052]The case where the case where they are tens of K bytes – hundreds of K bytes and the still picture data of several sheets whose still picture data size of one sheet is comparatively large are used as user data as the third example is explained using drawing 13. In this caseif AAU of the composition of the non-sound data shown in drawing 11 and drawing 12and user datasince user data size is largeAAU of the composition of silent data and user data will increase. Therefore since a non-sound will continue for a long time if this music stream is decodedit has composition of the music stream shown in drawing 13 in this case.

[0053]As for user data and 1303in drawing 131301 is [ user data and 1305 ] music data music data and 1304 music data and 1302.

[0054]Although a silent portion can be lost with constituting AAU from the music data 1301 and the user data 1302it is necessary to repress music data being conscious of the size beforehand used as user data in this case. The operation at this time is explained using drawing 2.

[0055]In drawing 2the music stream read from the recording medium 201 is sent to the music decoder 203 by the switch 202. Thenbeing conscious of the size beforehand used as user datamusic data is repressed with the music encoder 204. In the still picture adjunct 206the still picture data read from the recording medium 201 so that it might become the composition shown in drawing 13 at the music stream generated with the music encoder 204 is addedand it records on the recording medium 201 after that.

[0056]The case where take music and a synchronization and textssuch as wordsare displayed on a screen as the fourth example is explained using drawing 2 and drawing 14.

[0057]In drawing 2the music stream read from the recording medium 201 is sent to the music decoder 203 by the switch 202. Thenbeing conscious of the size beforehand used as user datamusic data is repressed with the music encoder 204. On the other handthe output of the music decoder 203 is inputted into the voice recognition part 205and is added to the music stream which has recognized \*\*\*\*\*outputted the text data and was outputted by the synchronous text adjunct 208 from the music encoder 204 at any time. Thenit records on the recording medium 201.

[0058]Next the music stream which the synchronous text adjunct 208 generates is explained.

[0059]As for music data and 1404 in drawing 14 music data and 1402 are [ music data and 1406 ] user data user data and 1405 user data and 1403 1401.

[0060]Thus it realizes by considering all the AAU(s) as the composition of music data and user data.

[0061]How to take out user data is explained using drawing 16. In drawing 16 the music stream with user data read from the recording medium 1601 is decoded by the music decoder 1602 and is outputted from the music output terminal 1609. On the other hand in the user data judgment part 1603 the data from data after music decoding was carried out to the header of the next AAU is investigated. At the time of 0 or 1 no user data are altogether outputted from the generating picture terminal 1608 to the next AAU as what not existing. When the header of still picture data such as JPEG and GIF is not found by the next AAU user data is sent to the text display area 1606 as text data and is outputted from the generating picture terminal 1608 at any time.

[0062]

[Effect of the Invention] By embedding in this invention at music data without making user data such as a track name a singer name words CD jackets and a singer's still picture data different from music data The user data additional equipment which simplify management with music data and user data and it is made to synchronize with music and user data is processed and can also be displayed on a monitor etc. at any time and user data playback equipment are provided.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is a figure showing the first example of the user data additional equipment of this invention.

[Drawing 2] It is a figure showing the second example of the user data additional equipment of this invention.

[Drawing 3] It is a figure showing signs that music data is downloaded using a communications satellite.

[Drawing 4] It is a figure showing signs that music data is downloaded from a network.

[Drawing 5] It is a figure showing the first example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 1.

[Drawing 6] It is a figure showing the second example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 1.

[Drawing 7] It is a figure showing the third example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 1.

[Drawing 8] It is a figure showing the fourth example of the lineblock diagram of the

music stream generated with the user data additional equipment of drawing 1.

[Drawing 9]It is a figure showing an example of the first example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 2.

[Drawing 10]It is a figure showing other examples of the first example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 2.

[Drawing 11]It is a figure showing an example of the second example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 2.

[Drawing 12]It is a figure showing other examples of the second example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 2.

[Drawing 13]It is a figure showing the third example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 2.

[Drawing 14]It is a figure showing the fourth example of the lineblock diagram of the music stream generated with the user data additional equipment of drawing 2.

[Drawing 15]It is a figure explaining the composition of the conventional music stream.

[Drawing 16]It is a figure showing the example of the user data playback equipment of this invention.

[Description of Notations]

101 ... Image input terminal

102 ... Music input terminal

103 ... Text input terminal

104 ... Still picture encoder

105 ... Music encoder

106 ... Switch

107 ... Voice recognition part

108 ... Still picture adjunct

109 ... Text adjunct

110 ... Synchronous text adjunct

111 ... User data additional equipment

112 ... Recording medium

---

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2000-305588  
(P2000-305588A)

(43)公開日 平成12年11月2日(2000.11.2)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	テマコード <sup>*</sup> (参考)
G 1 0 L 15/00		G 1 0 L 3/00	5 5 1 G 5 C 0 5 3
13/00			Q 5 D 0 1 5
15/10			5 3 1 N 5 D 0 4 5
15/28			5 5 1 P
19/02		7/04	G
審査請求 未請求 請求項の数24 O L (全 13 頁) 最終頁に続く			

(21)出願番号 特願平11-115981

(22)出願日 平成11年4月23日(1999.4.23)

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 今井 勉

神奈川県横浜市戸塚区吉田町292番地株式

会社日立製作所マルチメディアシステム開

発本部内

(72)発明者 尾鷲 仁朗

神奈川県横浜市戸塚区吉田町292番地株式

会社日立製作所マルチメディアシステム開

発本部内

(74)代理人 100075096

弁理士 作田 康夫

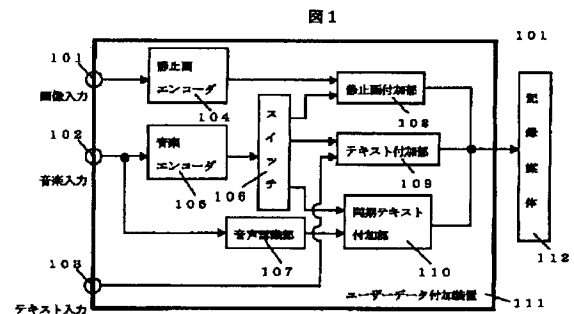
最終頁に続く

(54)【発明の名称】 ユーザーデータ付加装置及びユーザーデータ再生装置

(57)【要約】

【課題】音楽データとユーザーデータとの管理を簡単にし、また、音楽と同期させてユーザーデータを処理して随時モニタ等に表示することも可能なユーザーデータ付加装置及びユーザーデータ再生装置を提供する。

【解決手段】本発明では、曲名、歌手名、歌詞、CDジャケットや歌手の静止画データといったユーザーデータを音楽データと別にすることなく音楽データに埋め込むユーザーデータ付加装置及びユーザーデータ再生装置を提供する。これにより、音楽データを再生中に静止画や文字データを表示したり、音楽と同期させてユーザーデータを連続的に表示することも可能となる。



**【特許請求の範囲】**

【請求項1】画像を入力する画像入力端子と、音楽を入力する音楽入力端子と、テキストを入力するテキスト入力端子と、前記入力された画像を圧縮する静止画エンコーダと、前記入力された音楽を圧縮する音楽エンコーダと、前記圧縮された音楽ストリームを静止画付加部またはテキスト付加部または同期テキスト付加部に振り分けるスイッチと、前記入力された音楽から音声認識してテキストデータにする音声認識部と、前記圧縮された音楽ストリームに圧縮された静止画データを付加する静止画付加部と、前記圧縮された音楽ストリームにテキスト入力端子から入力されたテキストデータを付加するテキスト付加部と、前記圧縮された音楽ストリームに音声認識部でテキスト化されたデータを付加する同期テキスト付加部とを備えたことを特徴とするユーザーデータ付加装置。

【請求項2】圧縮された音楽ストリームを静止画付加部またはテキスト付加部または同期テキスト付加部に振り分けるスイッチと、前記圧縮された音楽ストリームを伸長する音楽デコーダと、前記伸長された音楽データを再圧縮する音楽エンコーダ、前記伸長された音楽データから音声認識してテキストデータにする音声認識部と、前記圧縮された音楽ストリームに圧縮された静止画データを付加する静止画付加部と、前記圧縮された音楽ストリームにテキストデータを付加するテキスト付加部と、前記圧縮された音楽ストリームに音声認識部でテキスト化されたデータを付加する同期テキスト付加部とを備えたことを特徴とするユーザーデータ付加装置。

【請求項3】圧縮された音楽ストリームを伸長する音楽デコーダと、前記伸長された音楽ストリームの音楽データ以外のデータが存在するのかを判定するユーザーデータ判定部と、ユーザーデータを静止画デコーダ又はテキスト表示部に送るスイッチと、前記圧縮された静止画データを伸長する前記静止画デコーダと、テキストのユーザーデータを処理する前記テキスト表示部と、前記静止画像や前記テキストを出力する画像出力端子と、前記音楽データを出力する音楽出力端子とを備えたことを特徴とするユーザーデータ再生装置。

【請求項4】請求項1記載において、前記静止画付加部は、音楽ストリームの生成時に、ユーザーデータとして静止画データを埋め込んで生成することを特徴とするユーザーデータ付加装置。

【請求項5】請求項1記載において、前記テキスト付加部は、音楽ストリームの生成時に、ユーザーデータとして文字データが埋め込んで生成することを特徴とするユーザーデータ付加装置。

【請求項6】請求項1記載において、前記同期テキスト付加部は、音楽ストリームの生成時に、ユーザーデータとして音楽に同期した文字データが埋め込んで生成することを特徴とするユーザーデータ付加装置。

【請求項7】請求項1記載において、前記静止画付加部またはテキスト付加部は、音楽ストリームの曲の先頭のAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が、無音データとユーザーデータで構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項8】請求項1記載において、前記静止画付加部またはテキスト付加部は、音楽ストリームの曲の先頭から数個のAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が、無音データとユーザーデータから構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項9】請求項1記載において、前記静止画付加部またはテキスト付加部は、音楽ストリームの数個のAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が、音楽データとユーザーデータから構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項10】請求項1記載において、前記同期テキスト付加部は、音楽ストリームのすべてのAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が音楽データとユーザーデータから構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項11】請求項1記載において、前記音楽エンコーダは、1AAU(Audio Access Unit、オーディオ・アクセス・ユニット)の中のあらかじめユーザーデータの領域を残して圧縮処理を行うことを特徴とするユーザーデータ付加装置。

【請求項12】請求項2記載において、前記静止画付加部は、音楽ストリームの生成時に、ユーザーデータとして静止画データを埋め込んで生成することを特徴とするユーザーデータ付加装置。

【請求項13】請求項2記載において、前記同期テキスト付加部は、音楽ストリームの生成時に、ユーザーデータとして音楽に同期した文字データが埋め込んで生成することを特徴とするユーザーデータ付加装置。

【請求項14】請求項2記載において、前記同期テキスト付加部は、音楽ストリームの生成時に、ユーザーデータとして音楽に同期した文字データを埋め込んで生成することを特徴とするユーザーデータ付加装置。

【請求項15】請求項2記載において、前記静止画付加部またはテキスト付加部は、音楽ストリームの曲の先頭のAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が、無音データとユーザーデータで構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項16】請求項2記載において、前記静止画付加部またはテキスト付加部は、音楽ストリームの曲の先頭から数個のAAU(Audio Access Unit、オーディオ・

アクセス・ユニット)が、無音データとユーザーデータから構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項17】請求項2記載において、前記静止画付加部またはテキスト付加部は、音楽ストリームの数個のAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が、音楽データとユーザーデータから構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項18】請求項2記載において、前記同期テキスト付加部は、音楽ストリームのすべてのAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が、音楽データとユーザーデータから構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項19】請求項2記載において、前記静止画付加部またはテキスト付加部は、音楽ストリームの曲の最後のAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が、無音データとユーザーデータから構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項20】請求項2記載において、前記静止画付加部またはテキスト付加部は、音楽ストリームの曲の最後から数個のAAU(Audio Access Unit、オーディオ・アクセス・ユニット)が、無音データとユーザーデータから構成された音楽ストリームを生成することを特徴とするユーザーデータ付加装置。

【請求項21】請求項2記載において、前記音楽エンコーダは、AAU(Audio Access Unit、オーディオ・アクセス・ユニット)の中あらかじめユーザーデータの領域を残して圧縮処理を行うことを特徴とするユーザーデータ付加装置。

【請求項22】請求項3記載において、前記ユーザーデータ判定部は、無音データもしくは音楽データに続く次のAAU(Audio Access Unit、オーディオ・アクセス・ユニット)までのデータがすべて0またはすべて1以外の場合にはユーザーデータとしてみなして処理を行うことを特徴とするユーザーデータ再生装置。

【請求項23】請求項3記載において、前記ユーザーデータ判定部は、AAU(Audio Access Unit、オーディオ・アクセス・ユニット)が無音データの場合には次のAAUまでのデータをユーザーデータとしてみなして処理を行うことを特徴とするユーザーデータ再生装置。

【請求項24】請求項22または請求項23記載において、前記ユーザーデータ判定部は、ユーザーデータ中の圧縮された静止画のヘッダを検索し、ヘッダが見つかったときにはユーザーデータを圧縮された静止画データと判定し、ヘッダが見つからないときにはテキストデータと判定することを特徴とするユーザーデータ再生装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音楽をデジタル圧縮して記録媒体に記録する装置及び通信衛星を使用した音楽配信やパソコン等を使用したネットワークの音楽配信でダウンロードされた音楽ストリームを再生する装置に係わり、特に圧縮された音楽ストリームにテキストデータや静止画データを付加するユーザーデータ付加装置及びユーザーデータ再生装置に関する。

【0002】

【従来の技術】従来のデジタル圧縮をして音楽配信を行う例として、通信衛星を使用した音楽配信やパソコンを使用しインターネット上のホームページからダウンロードする音楽配信がある。

【0003】通信衛星を使い音楽配信を行うものとして、PC用データ放送である「スカイ・パーフェクトPC! (SKY Perfect PC!)」がある。これは、通信衛星を使用してホームページのデータ、プログラム、音楽データのダウンロード等を行う。

【0004】一方、パソコン等を使用したインターネットを使用した音楽配信を行うものとしてはデジタル圧縮されたWeb上の音楽データを一定の料金を払いダウンロードし、その音楽データを伸長することのできるパソコン等で再生する。

【0005】また、圧縮されていない音楽をリアルタイムにデジタル圧縮する音楽圧縮装置がある。

【0006】

【発明が解決しようとする課題】従来の音楽配信では、音楽ストリームのみのダウンロードであり、ハードディスクやメモリ等の記録媒体に記録されたファイル名にユーザーデータである曲名や歌手名等を付けていた。音楽データと別にユーザーデータとして曲名や歌手名、静止画データ等の情報もダウンロードできる場合もあるが、その場合にはデジタル圧縮された音楽データとユーザーデータをファイル名を同じにしたり同じディレクトリーに記録したりする必要がある。しかしながら、音楽データとユーザーデータとを同期を取って処理をする、例えば音楽に合わせて歌詞を随時表示したりすることは困難であった。

【0007】また、リアルタイムに音楽を録音する場合には音楽のみでありユーザーデータは無かった。

【0008】

【課題を解決するための手段】上記課題を解決するため、本発明では、曲名、歌手名、歌詞、CDジャケットや歌手の静止画データといったユーザーデータを音楽データと別にすることなく音楽データに埋め込む装置及びユーザーデータ再生装置を提供する。これにより、音楽データを再生中にユーザーデータを表示したり、音楽と同期させてユーザーデータを連続的に表示することが可能となる。

【0009】



【発明の実施の形態】以下、本発明の実施例を図面を用いて説明する。

【0010】ユーザーデータとして、曲名・歌手名、歌詞、CDジャケットや歌手の静止画データを付加する装置として音楽をリアルタイムに圧縮しながらユーザーデータを付加する場合と配信された音楽ストリームを一旦ハードディスクやメモリ等の記録媒体に記録した音楽ストリームにユーザーデータを付加する方法について説明する。

【0011】第一の実施例として、マイクや外部装置からの音楽をリアルタイムに圧縮しながらユーザーデータを付加する場合について説明する。音楽のデジタル圧縮方式としてMPEG Audio Layer IIの場合について図1、図16を用いて説明する。

【0012】図1において、101は画像入力端子、102は音楽入力端子、103はテキスト入力端子、104は静止画エンコーダ、105は音楽エンコーダ、106はスイッチ、107は音声認識部、108は静止画付加部、109はテキスト付加部、110は同期テキスト付加部、111はユーザーデータ付加装置、112はハードディスクやメモリ等の記録媒体である。

【0013】画像入力端子101にはカメラ等の撮像装置が接続され、音声入力端子102には、マイクもしくは音楽再生装置が接続され、テキスト入力端子103にはキーボード等と接続される。音楽入力端子102から取り込まれた音楽は音楽エンコーダ105によりデジタル圧縮され音楽ストリームを生成する。次に、付加するユーザーデータに応じてスイッチ106により静止画付加部108、テキスト付加部109、同期テキスト付加部110に切り替える。静止画付加部108では、音楽エンコーダ105が生成した音楽ストリームに画像入力端子101で入力された画像を静止画エンコーダ104によってJPEG (Joint Photographic Coding Experts Group)、GIF (Graphics Interchange Format)等のフォーマットに圧縮した静止画画像を付加する。テキスト付加部109では、音楽エンコーダ105が生成した音楽ストリームにテキスト入力端子103からキーボード等で入力されたテキストデータを付加する。同期テキスト付加部110では、音楽エンコーダ105が生成した音楽ストリームに音楽入力端子102から入力された音楽データから音声認識部107で音声情報から歌詞を認識しテキスト化して付加する。

【0014】ここで、音楽エンコーダ105が生成する音楽ストリームについて図15を用いて説明する。

【0015】図15において、1501は音楽ストリームを構成しているAAU (Audio Access Unit、オーディオ・アクセス・ユニット) で一つ一つ単独で音楽信号に復号できる最小単位である。AAU 1501は、ヘッダ1502、ビットアロケーション1503、SCFS I (Scalefactor Selection Information、スケールフ

クタ選択情報) 1504、スケールファクタ1505、サンプル1506で構成している。

【0016】第一の具体例としてデータ量の少ない曲名・歌手名のテキストをユーザーデータとして付加する場合のテキスト付加部108が生成する音楽ストリームを図5を用いて説明する。

【0017】図5において、501は無音データ、502はユーザーデータ、503は音楽データ、504は音楽データ、505は無音データ、506ユーザーデータである。501の無音データから504の音楽データまでが一曲分の音楽ストリームであり、505の無音データ以降が次の曲の音楽ストリームを示している。この場合には、各音楽ストリームを構成する複数のAAUの内の先頭のAAUを、無音データ501とユーザーデータ502から構成し即ち無音データ501にユーザーデータ502を付加する場合が示してある。

【0018】ここで、無音データ501の作成の仕方を説明する。MPEG Audio、Layer II、サンプリング周波数48kHz、ビットレート256kbps、ステレオの場合では、1AAUは6144ビットである。無音データは、ヘッダ1602の32ビットを0xFFFD C40Cとしビットアロケーション1603の88ビットをすべて0とすることにより、ヘッダ(32ビット)とビットアロケーション(88ビット)の120ビットだけで無音データとすることができる。つまり、1AAUの残り6144-120=6024ビットをユーザーデータに割り当てることが可能となる。

【0019】次に、ユーザーデータの取り出し方について図16を用いて説明する。図16において、1601は記録媒体、1602は音楽デコーダ、1603はユーザーデータ判定部、1604はスイッチ、1605は静止画デコーダ、1606はテキスト表示部、1607はユーザーデータ再生装置、1608は画像出力端子、1609は音楽出力端子である。

【0020】記録媒体1601から読み出されたユーザーデータ付き音楽データストリームは、音楽デコーダ1602でデコードされ、音楽出力端子1609に接続されているスピーカから出力される。

【0021】ユーザーデータ判定部1603では音楽デコードされた後のデータから次のAAUのヘッダまでのデータを調べる。通常、エンコーダは1AAUをほとんど埋めつくすように計算しながら音楽データを圧縮し、余った部分にはすべて0や1を埋め込む。そのため、次のAAUまですべて0や1のときにはユーザーデータは存在しないものとして画像出力端子1608からは何も出力しない。次のAAUまでにJPEGやGIF等の静止画データのヘッダが見つからなかったときにはテキストデータと見なしてテキスト表示部1606でテキストを解読し画像出力1608に接続されているモニタ等にテキストを表示する。また、通常は全く無音のAAUは

存在しないので無音のAAUのときにはユーザーデータが含まれているAAUとして処理を行ってもよい。

【0022】一方、ユーザーデータ再生装置1607以外のユーザーデータを処理する機能のない再生装置で無音データとユーザーデータから構成されているAAUをデコードするしたと仮定する。この場合には、無音データのみを処理しそれに続くユーザーデータは無視されることになる。そのため、スピーカーからは音が聞こえないで無音が再生されることになるためなら影響を及ぼさない。1AAUの再生時間は、1AAUに含まれているサンプルは1152サンプルであるので、 $1/48$  (kHz)  $\times 1152 = 24$  (ms) となる。従って、図1の音楽ストリームの構成とした場合には曲と曲のつながりで24ms無音データとなる。

【0023】第二の具体例として、ユーザーデータとしてJPEG (Joint Photographic Coding Experts Group)、GIF (Graphics Interchange Format) 等で圧縮したCDジャケットや歌手の1枚分の静止画データとし、そのサイズが比較的小さい場合の静止画付加部が生成する音楽ストリームを図6を用いて説明する。

【0024】図6において、601は無音データ、602はユーザーデータ、603は無音データ、604はユーザーデータ、605は音楽データ、606は音楽データである。601の無音データから606の音楽データまでが一曲分の音楽データである。

【0025】JPEGやGIF等で圧縮した比較的小さい静止画データでも数kバイト～数十kバイトのユーザーデータサイズを必要とする。そのため、無音データとユーザーデータから構成されるAAUが1つでは足りないため静止画データ量に合うように数個用意する。無音データとユーザーデータの構成では上記で説明したようにユーザーデータに6024ビット使用可能である。例えば、1kバイト ( $1024 \times 8 = 8192$  ビット) の静止画データをユーザーデータとして使用するときには、無音データとユーザーデータの構成のAAUを2個とすることによりこれに対応する。

【0026】ユーザーデータの取り出し方について図16を用いて説明する。

【0027】図16において、記録媒体1601から読み出されたユーザーデータ付き音楽データストリームは、音楽デコーダ1602でデコードされ、音楽出力端子1609に出力される。上記具体例1の場合と同様に、ユーザーデータ判定部1603では音楽デコードされた後のデータから次のAAUのヘッダまでのデータを調べる。次のAAUまですべて0や1のときにはユーザーデータは存在しないものとして画像出力端子1608からは何も出力しない。次のAAUまでにJPEGやGIF等の静止画データのヘッダが見つかったときにはスイッチ1604により静止画デコーダ1605の処理を行い画像出力に接続されているモニタ等に静止画を出力

するのだが、静止画データが1AAUに収まっていない場合は次のAAUにも残りのデータがあるものとして静止画デコーダ処理を行い画像出力端子から出力する。

【0028】第三の具体例として、1枚の静止画データサイズが比較的大きい数十kバイト～数百kバイトの場合や数枚の静止画データをユーザーデータとして使用する場合を図1、図7を用いて説明する。この場合には、図6で示したAAUの構成のように無音データとユーザーデータの構成とするとユーザーデータサイズが大きいため無音データとユーザーデータの構成のAAUが多くなってしまう。そのため、この音楽ストリームをデコードすると無音が長く。そこで、ユーザーデータサイズが大きい場合には図7に示す音楽ストリームの構成とする。

【0029】図7において、701は音楽データ、702はユーザーデータ、703は音楽データ、704はユーザーデータ、705は音楽データ、706はユーザーデータ、707は音楽データである。701の音楽データから707の音楽データまでが一曲分の音楽データである。このように、音楽データとユーザーデータでAAUを構成することで無音の部分を無くすることができる。しかし、この場合にはあらかじめユーザーデータとして使用するサイズを意識して音楽エンコーダ105で音楽データの圧縮を行う必要がある。

【0030】例えば、10kバイト ( $10 \times 1024 \times 8 = 81920$  ビット) の静止画データをユーザーデータとする場合に、1AAUの6144ビット中6044ビットを音楽データ701に割り当て残りの100ビットをユーザーデータ702に割り当てる。このような割り当てにすると、 $81920 / 100 = 819.2$  となり820個のAAUを音楽データ701とユーザーデータ702の構成とすることにより実現できる。1AAUは24msであるため  $820 \times 24 = 19680$  ms 分の音楽ストリームを音楽データとユーザーデータの構成とする。

【0031】ユーザーデータの取り出し方は、上記具体例2と同じであるので省略する。

【0032】第四の具体例として、音楽と同期を取り画面に歌詞を表示する場合を図1、図8を用いて説明する。

【0033】図8において、801は音楽データ、802はユーザーデータ、803は音楽データ、804はユーザーデータ、805は音楽データ、806はユーザーデータ、807は音楽データ、808はユーザーデータである。801の音楽データから806のユーザーデータまでが一曲分の音楽ストリームであり、807の音楽データ以降が次の曲の音楽データを示している。

【0034】図1において、音楽入力端子102から入力された音楽は音声認識部107で音声を認識しテキストデータとする。同期テキスト付加部110では音楽エ

ンコーダ105から出力された音楽ストリームにテキストデータを随時付加する。具体例3と同様に音楽エンコーダ105ではあらかじめユーザーデータの部分を空けた音楽ストリームを作成する。

【0035】ユーザーデータの取り出し方について図16を用いて説明する。図16において、記録媒体1601より読み出されたユーザーデータ付きの音楽ストリームは音楽デコーダ1602でデコードされ音楽出力端子1609から出力される。一方、ユーザーデータ判定部1603では音楽デコードされた後のデータから次のAAUのヘッダまでのデータを調べる。次のAAUまですべて0や1のときにはユーザーデータは存在しないものとして画像出力端子1608からは何も出力しない。次のAAUまでにJPEGやGIF等の静止画データのヘッダが見つからなかったときにはユーザーデータはテキストデータとしてテキスト表示部1606に送られ随時画像出力部1608から出力する。

【0036】第二の実施例として、通信衛星やネットワークを使い配信された音楽ストリームを一旦ハードディスクやメモリ等の記録媒体に記録した音楽ストリームにユーザーデータを付加する方法について説明する。記録媒体に音楽データをダウンロードする方法を図3、図4を用いて説明する。

【0037】図3は通信衛星を使用してダウンロードする様子を説明する図である。図3において、301はパラボラアンテナ、302は受信ボード、303は記録媒体、304はパソコンである。通信衛星から送られた数チャンネルの音楽ストリームはパラボラアンテナ301で受信され受信ボード302により料金を払った音楽ストリームやサンプルの音楽ストリームのみ記録媒体303に記録される。その後、ダウンロードした音楽ストリームをパソコン等で伸長し再生するものである。

【0038】図4はインターネット上の情報を電話回線等を介してダウンロードする様子を説明する図である。図4において、401はインターネット、402はモデム、403は記録媒体404はパソコンを示している。この場合にはWebに音楽データがファイルとして登録されており一定の料金を払い電話回線を使ってモデム402を介して記録媒体403にダウンロードする。その後、ダウンロードした音楽ストリームををパソコン等で伸長し再生するものである。

【0039】まず、図2を用いて記録媒体にダウンロードした音楽ストリームにユーザーデータを付加する動作について説明する。図2において、201は記録媒体、202はスイッチ、203は音楽デコーダ、204は音楽エンコーダ、205は音声認識部、206は静止画付加部、207はテキスト付加部、208は同期テキスト付加部、209はユーザーデータ付加装置である。

【0040】通信衛星やネットワークを使用した音楽配信でのデジタル圧縮ではMPEG Audio、Layer

IIIより高圧縮であるMPEG Audio、Layer III (MP3) が多く使われている。

【0041】第一の実施例と同様にユーザーデータとしては、曲名・歌手名、歌詞、CDジャケットや歌手の静止画データが考えられる。それぞれの場合について本発明を用いて生成する音楽ストリームを説明する。

【0042】第一の具体例としてデータ量の少ない曲名・歌手名のテキストデータをユーザーデータとして使用する場合を図9、図10を用いて説明する。

【0043】図9において、901は無音データ、902はユーザーデータ、903は音楽データ、904は音楽データである。

【0044】図2において、記録媒体201から読み出された音楽ストリームはスイッチ202によりテキスト付加部207に送られテキスト付加部207でテキストデータを付加され記録媒体201に記録される。

【0045】データ量の少ない曲名・歌手名のテキストデータをユーザーデータとする場合には、図9に示すように、音楽ストリームを構成するAAUの内の先頭のAAUを、無音データ901とユーザーデータ902から構成し即ち無音データ901にユーザーデータ902を付加することにより実現する。

【0046】一方、ダウンロードされた音声ストリームは曲毎に1ファイルとなっているためファイルの最後(曲の最後)のAAUが検索しやすい。そのため、図10に示すようにファイルの最後にユーザーデータを含んだAAU(無音データ1003、ユーザーデータ1004)を付加することでも実現できる。

【0047】ユーザーデータの取り出し方について図16を用いて説明する。

【0048】図16において、記録媒体1601から読み出されたユーザーデータ付き音楽データストリームは、音楽デコーダ1602でデコードされ、音楽出力端子1609から出力される。ユーザーデータ判定部1603では音楽デコードされた後のデータから次のAAUのヘッダまでのデータを調べる。次のAAUまですべて0や1のときにはユーザーデータは存在しないものとして画像出力端子1608からは何も出力しない。次のAAUまでにJPEGやGIF等の静止画データのヘッダが見つからなかったときにはテキストデータと判断してテキスト表示部1606で処理を行い1608から出力する。

【0049】第二の具体例として、ユーザーデータとしてJPEG、GIF等で圧縮したCDジャケットや歌手の1枚分の静止画データとし、そのサイズが比較的小さい場合を図11、図12を用いて説明する。

【0050】図11において、1101は無音データ、1102はユーザーデータ、1103は無音データ、1104はユーザーデータ、1105は音楽データである。

【0051】圧縮した静止画データサイズが小さい場合でも数kバイト〜数十kバイトのユーザーデータサイズを必要とする。そのため、無音データとユーザーデータから構成されるAAUが1つでは足りないため静止画データ量に合うように数個をファイルの先頭に付加することにより実現する。また、図12に示すようにユーザーデータを含んだAAU（無音データ1204、ユーザーデータ1205）を最後に数個付加した構成とすることも実現できる。

【0052】第三の具体例として、1枚の静止画データサイズが比較的大きい数十kバイト〜数百kバイトの場合や数枚の静止画データをユーザーデータとして使用する場合は図13を用いて説明する。この場合には図11、図12に示した無音データとユーザーデータの構成のAAUとするとユーザーデータサイズが大きいため無音データとユーザーデータの構成のAAUが多くなってしまふ。そのためこの音楽ストリームをデコードすると無音が長く続いてしまうため、この場合は図13に示す音楽ストリームの構成とする。

【0053】図13において、1301は音楽データ、1302はユーザーデータ、1303は音楽データ、1304はユーザーデータ、1305は音楽データである。

【0054】音楽データ1301とユーザーデータ1302でAAUを構成することで無音の部分無くすることができるが、この場合にはあらかじめユーザーデータとして使用するサイズを意識して音楽データの再圧縮を行う必要がある。このときの動作を図2を用いて説明する。

【0055】図2において、記録媒体201から読み込まれた音楽ストリームはスイッチ202により音楽デコーダ203に送られる。その後、音楽エンコーダ204であらかじめユーザーデータとして使用するサイズを意識して音楽データの再圧縮を行う。静止画付加部206では音楽エンコーダ204で生成された音楽ストリームに図13に示した構成となるように記録媒体201から読み出された静止画データを付加しその後記録媒体201に記録する。

【0056】第四の具体例として、音楽と同期を取って画面に歌詞等のテキストを表示する場合を図2、図14を用いて説明する。

【0057】図2において、記録媒体201から読み出された音楽ストリームはスイッチ202により音楽デコーダ203に送られる。その後、音楽エンコーダ204であらかじめユーザーデータとして使用するサイズを意識して音楽データの再圧縮を行う。一方、音楽デコーダ203の出力は音声認識部205に入力され、音声声を認識しテキストデータを出力し同期テキスト付加部208により音楽エンコーダ204から出力された音楽ストリームに随時付加する。その後、記録媒体201に記録

する。

【0058】次に、同期テキスト付加部208が生成する音楽ストリームについて説明する。

【0059】図14において、1401は音楽データ、1402はユーザーデータ、1403は音楽データ、1404はユーザーデータ、1405は音楽データ、1406はユーザーデータである。

【0060】このように、すべてのAAUを音楽データとユーザーデータの構成とすることにより実現する。

【0061】ユーザーデータの取り出し方について図16を用いて説明する。図16において、記録媒体1601より読み出されたユーザーデータ付きの音楽ストリームは音楽デコーダ1602でデコードされ音楽出力端子1609から出力される。一方、ユーザーデータ判定部1603では音楽デコードされた後のデータから次のAAUのヘッダまでのデータを調べる。次のAAUまですべて0や1のときにはユーザーデータは存在しないものとして画像出力端子1608からは何も出力しない。次のAAUまでにJPEGやGIF等の静止画データのヘッダが見つからなかったときにはユーザーデータはテキストデータとしてテキスト表示部1606に送られ随時画像出力端子1608から出力する。

【0062】

【発明の効果】本発明では、曲名、歌手名、歌詞、CDジャケットや歌手の静止画データといったユーザーデータを音楽データと別にすることなく音楽データに埋め込むことにより、音楽データとユーザーデータとの管理を簡単にし、また、音楽と同期させてユーザーデータを処理して随時モニタ等に表示することも可能なユーザーデータ付加装置及びユーザーデータ再生装置を提供する。

【図面の簡単な説明】

【図1】本発明のユーザーデータ付加装置の第一の実施例を示す図である。

【図2】本発明のユーザーデータ付加装置の第二の実施例を示す図である。

【図3】音楽データを通信衛星を用いてダウンロードする様子を示す図である。

【図4】音楽データをネットワーク上からダウンロードする様子を示す図である。

【図5】図1のユーザーデータ付加装置で生成される音楽ストリームの構成図の第一の具体例を示す図である。

【図6】図1のユーザーデータ付加装置で生成される音楽ストリームの構成図の第二の具体例を示す図である。

【図7】図1のユーザーデータ付加装置で生成される音楽ストリームの構成図の第三の具体例を示す図である。

【図8】図1のユーザーデータ付加装置で生成される音楽ストリームの構成図の第四の具体例を示す図である。

【図9】図2のユーザーデータ付加装置で生成される音楽ストリームの構成図の第一の具体例の一例を示す図である。

【図10】図2のユーザーデータ付加装置で生成される音楽ストリームの構成図の第一の具体例の他の例を示す図である。

【図11】図2のユーザーデータ付加装置で生成される音楽ストリームの構成図の第二の具体例の一例を示す図である。

【図12】図2のユーザーデータ付加装置で生成される音楽ストリームの構成図の第二の具体例の他の例を示す図である。

【図13】図2のユーザーデータ付加装置で生成される音楽ストリームの構成図の第三の具体例を示す図である。

【図14】図2のユーザーデータ付加装置で生成される音楽ストリームの構成図の第四の具体例を示す図である。

【図15】従来の音楽ストリームの構成を説明する図で

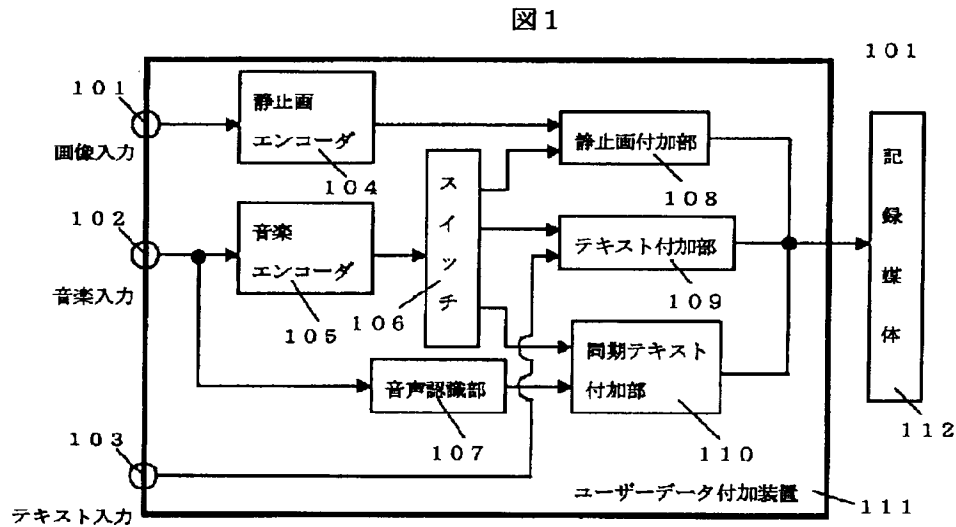
ある。

【図16】本発明のユーザーデータ再生装置の実施例を示す図である。

【符号の説明】

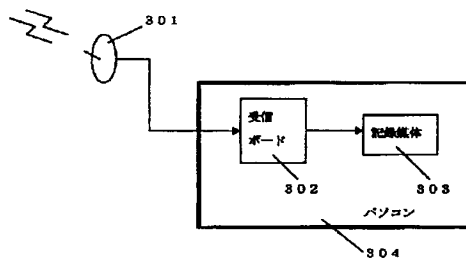
- 101・・・画像入力端子
- 102・・・音楽入力端子
- 103・・・テキスト入力端子
- 104・・・静止画エンコーダ
- 105・・・音楽エンコーダ
- 106・・・スイッチ
- 107・・・音声認識部
- 108・・・静止画付加部
- 109・・・テキスト付加部
- 110・・・同期テキスト付加部
- 111・・・ユーザーデータ付加装置
- 112・・・記録媒体

【図1】



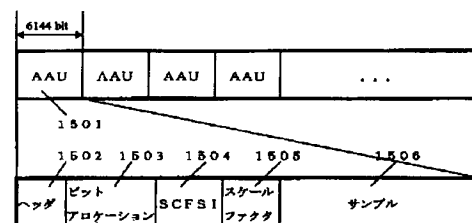
【図3】

図3



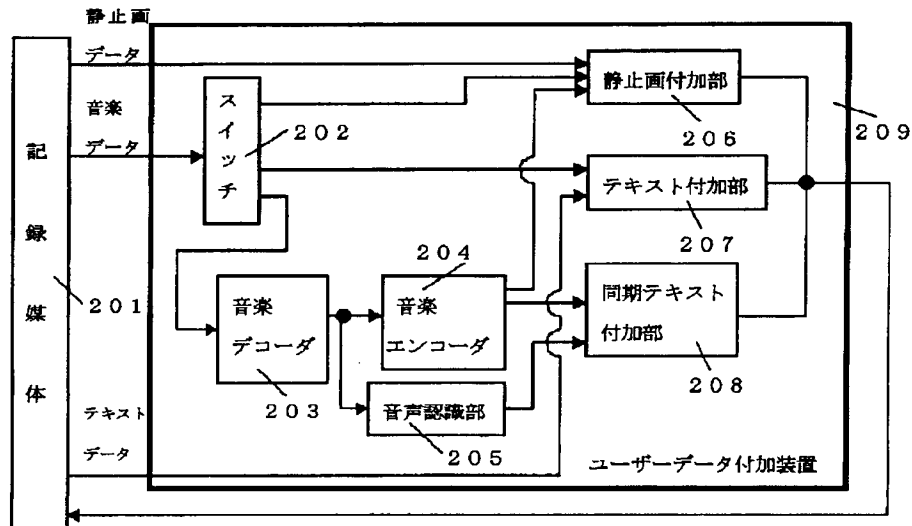
【図15】

図15



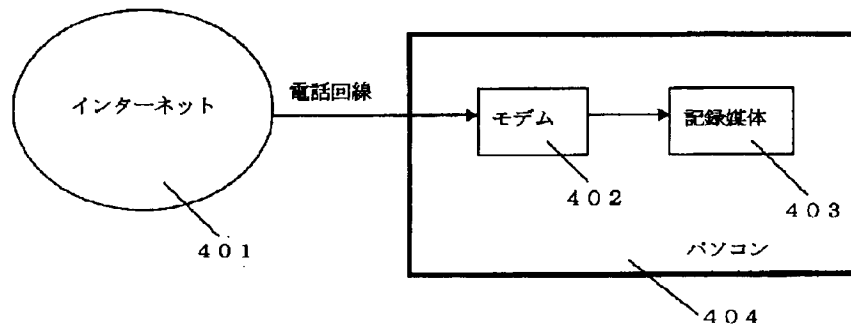
【図2】

図2



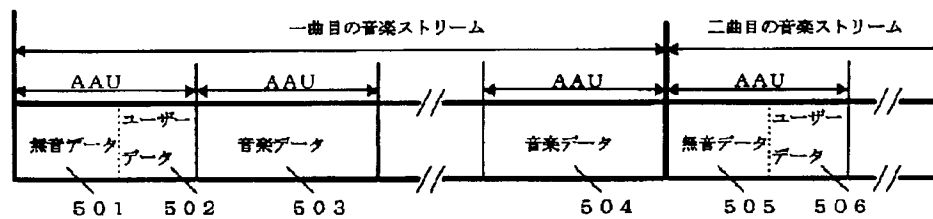
【図4】

図4



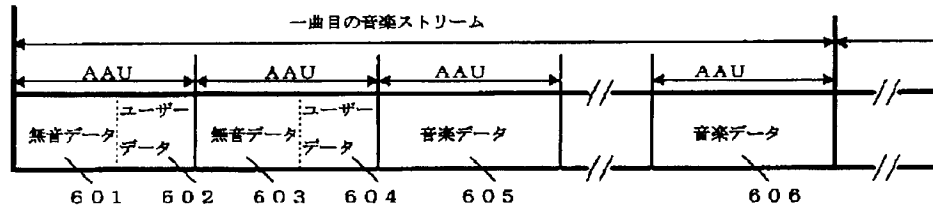
【図5】

図5



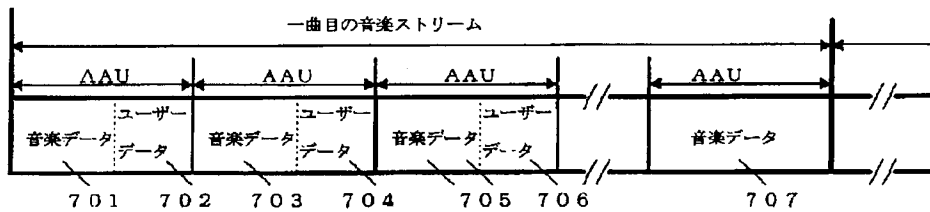
【図6】

図6



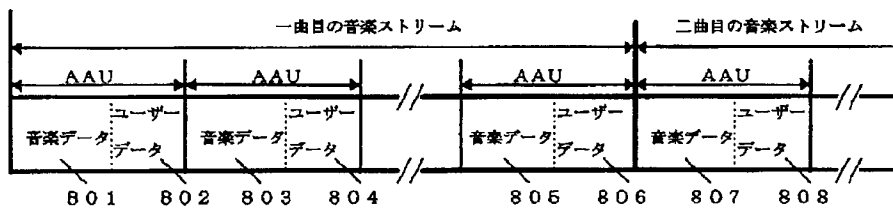
【図7】

図7



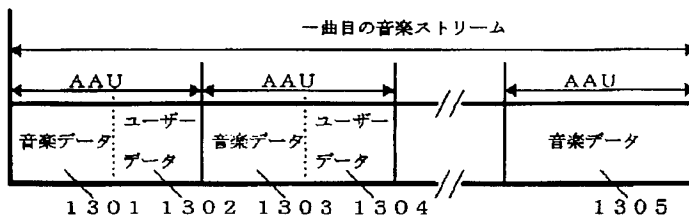
【図8】

図8



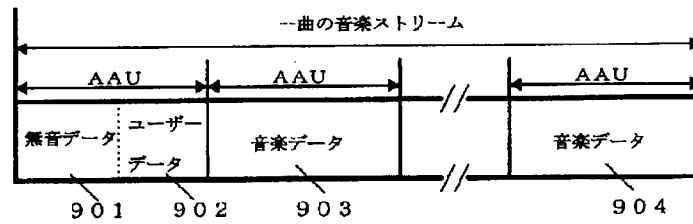
【図13】

図13



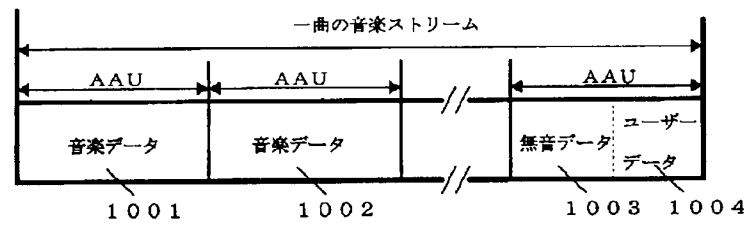
【図9】

図9



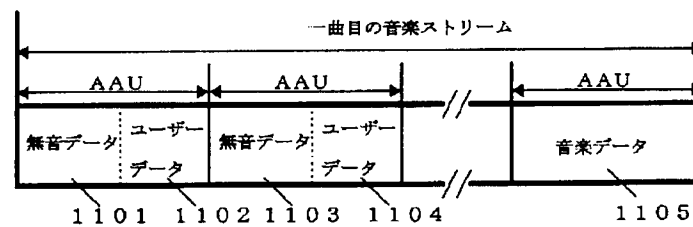
【図10】

図10



【図11】

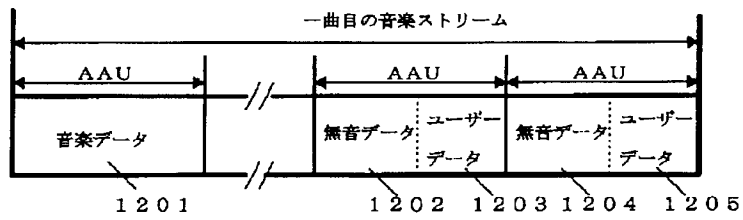
図11





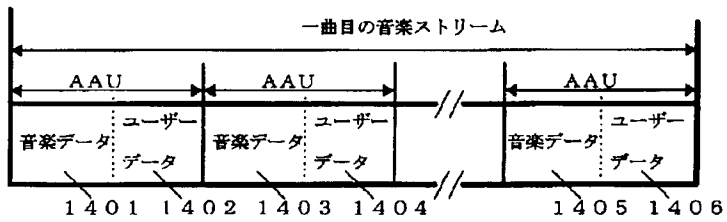
【図12】

図12



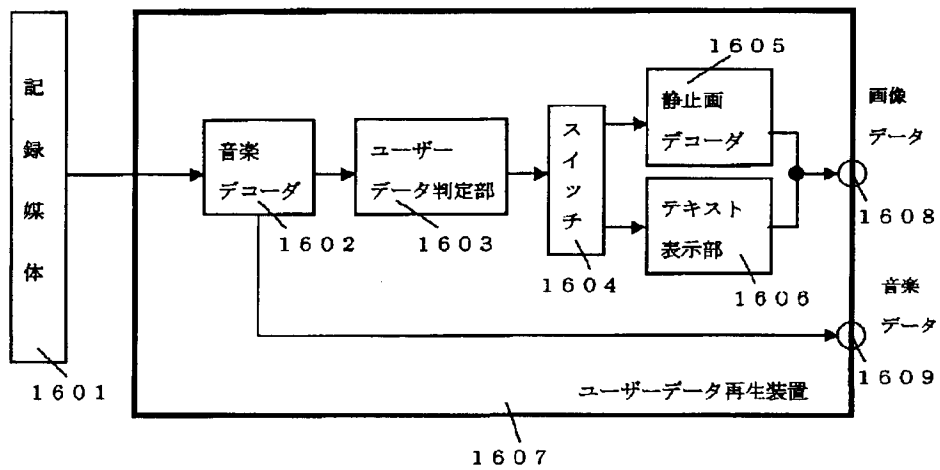
【図14】

図14



【図16】

図16



## フロントページの続き

(51) Int. Cl. 7	識別記号	F I	タームコード (参考)	
G 1 0 L	19/00	G 1 0 L	9/18	M
H 0 4 N	5/765	H 0 4 N	5/781	5 1 0 L
	5/781		5/91	R
	5/91			

F ターム (参考) 5C053 FA08 FA10 FA27 GB11 GB36  
LA02 LA11  
5D015 KK02  
5D045 AB01 DA20